



KONGERIKET NORGE  
The Kingdom of Norway

Bekreftelse på patentsøknad nr  
*Certification of patent application no*

1999 6338

Det bekreftes herved at vedheftede dokument er nøyaktig utskrift/kopi av ovennevnte søknad, som opprinnelig inngitt 1998.12.23

*It is hereby certified that the annexed document is a true copy of the above-mentioned application, as originally filed on 1998.12.23*

2003.07.17

*Freddy Strømmen*

Freddy Strømmen  
Seksjonsleder

*Line Reum*

Line Reum



PATENTSTYRET®  
Styret for det industrielle rettsvern

16

PATENTSTYRET

20.DES99 996338

20 DES. 1999

EK/KBN

17.12.99

E10515

UTSKILT FRA SØKNAD nr. 19986133 av 23/12-1998

Preben Lexow  
Fløensbakken 41A  
5009 Bergen

Oppfinner:

Søkeren

**Metoder for sekvensanalyse**

Foreliggende oppfinnelse er utskilt fra patentsøknad 19986133 som omfatter en metode for DNA-sekvensering som inneholder følgende trinn:

Første trinn tar utgangspunkt i en ren DNA-populasjon bestående av DNA-sekvensen som skal sekvenseres. DNA-molekylene kuttet/brekkes på en uspesifikk måte slik at det dannes en populasjon med DNA-molekyler bestående av biter (heretter kalt DNA-biter) av den opprinnelige sekvensen.

Annet trinn består i å erstatte baseparene i DNA-bitene med 4 ulike DNA-sekvenser (heretter kalt DNA-fragmenter) som representerer hver av de fire basene adenin, cytosin, guanin og tymin. Der hvor det har vært basepar A-T, settes det altså inn "fragment A", C-G byttes ut med "fragment C" osv. Dermed genereres nye DNA-molekyler hvor den opprinnelige baserekkefølgen på f.eks. ACGTT... erstattes med fragment A - fragment C - fragment G osv. Lengden på disse fire DNA-fragmentene kan i prinsippet variere i lengde fra 2bp til flere hundre kbp (eller mer om ønskelig), alt etter behov. Tilsvarende kan DNA-fragmentene inneholde reportergener og annen biologisk informasjon eller kun bestå av sekvenser uten kjent biologisk funksjon.

I tredje trinn avleses rekkefølgen av de fire typene DNA-fragmenter for hvert enkelt DNA-molekyl. Dermed finner man baserekkefølgen i de opprinnelige DNA-bitene indirekte.

I fjerde trinn benytter et dataprogram overlappen mellom DNA-bitene til å sette sammen informasjonen fra trinn 3 til sekvensen på DNA-sekvensene som ble brukt som utgangspunkt.

Foreliggende oppfinnelse omhandler sorteringsmekanismer til bruk ved sekvensanalyse.

## Presortering

I mange sammenhenger kan det være en fordel å innlede konverteringstrinnet med en presortering med utgangspunkt i de første baseparene i mål DNA'en. F.eks. kan man fordele mål DNA'en på 256 reagensrør. I det første røret gjennomfører man en prosedyre som plukker ut DNA biter som begynner med AAAA og gjør disse klare til et konverteringstrinn hvor de 4 neste baseparene konverteres. DNA molekylerne som ikke begynner med AAAA blir ikke klargjort og følgelig heller ikke konvertert i neste trinn. I rør nummer 2 gjør man det samme med DNA biter som begynner med AAAC, osv. I avlesningstrinnet vil man dermed få informasjon om 8 basepar: de fire første fordi man vet hvilket av 256 rørene konverteringen har foregått, mens de fire siste er konvertert og kan avleses direkte.

Det må særlig understrekes at presortering er særlig velegnet sammen med bruk av posisjonsmarkører, "optical mapping", o.l. Hvis hver signalkjede kan lokaliseres med en presisjon på noen få kb eller bedre trenger man som regel ikke å få informasjon om mer enn 8-10bp får å kunne rekonstruere sekvensen vellykket. Bruk av presortering gjør at man når dette målet uten å måtte resirkularisere DNA molekyler. Dermed unngår man bl.a. problemet med intermolekulære ligeringer som øker med størrelsen på DNA molekylerne.

## Presortering av 4 basepar på en DNA chips

Presorteringen kan naturligvis gjøres på svært mange måter. Som alternativ til å bruke 256 ulike rør kan man bruke en strategi hvor man har 256 ulike presorteringsadaptore fordelt på 256 ruter på en DNA chips. I rute 1 er det presorteringsadaptore med AAAA overheng, i rute 2 har de AAAC overheng, osv. Dermed vil DNA bitene sorteres slik at de med TTTT overheng festes til rute 1, GTTT overheng til rute 2, osv. Ved deretter å feste også den andre enden på DNA biten til underlaget, f.eks. med biotin/ streptavidin som illustrert i fig.32, kan man gå videre til neste konverteringstrinn uten at DNA molekylerne forlater posisjonen på avlesningsplaten. En annen strategi for å unngå at DNA molekylerne forlater sine posisjoner kan være å bruke en avlesningsplate som er inndelt i 256 brønner/rom.

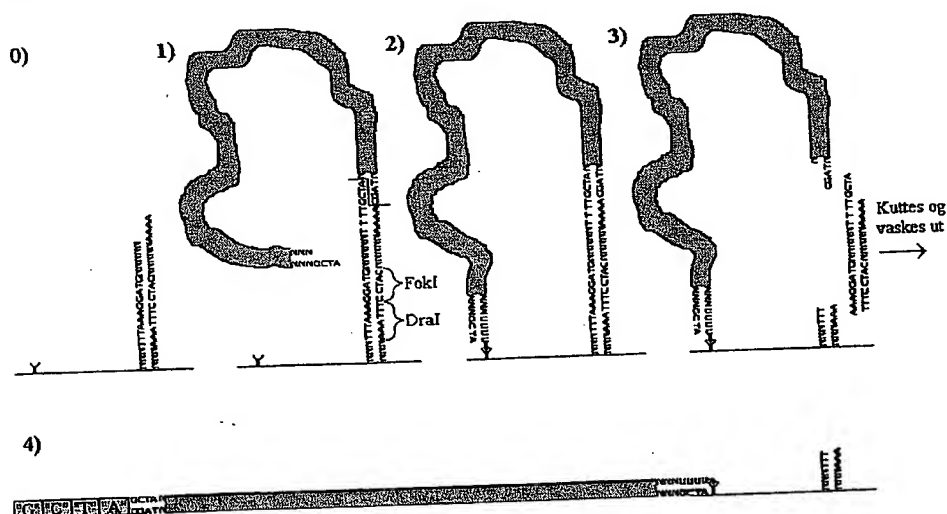


Fig.32. Eksempel på hvordan presorteringen kan gjøres på en DNA chips. 0) Utgangspunktet er en DNA chips som er oppdelt i 256 ruter. Hver rute inneholder presorteringsadaptore et overheng, et bindingssete for en klasse IIS restriksjons endonuklease og et bindingssete for en restriksjons endonuklease som lager et blunt end kutt. Overhengene varierer fra rute til rute slik at rute 1 har presorteringsadaptore med AAAA overheng, rute 2 har AAAC overheng, osv. Alle rutene er i tillegg dekket med et molekyl med bindingsegenskaper, f.eks. streptavidin. 1) Presorteringen innledes med at mål DNA'en kuttes i DNA biter og at endene på DNA bitene behandles slik at det dannes overheng på 4 baser. Disse kan deretter liggeres til presorteringsadapterene. DNA biter med TTTT-overheng vil ligere seg til rute 1 hvor presorteringsadaptore har det komplementære overhengen AAAA, DNA biter med GTTT-overheng vil ligere seg til rute 2,

osv. 2) Det andre overhenget på DNA biten behandles deretter slik at enden kan festes til underlaget. F.eks. kan man bruke klenow fylling for å merke enden med biotin, man kan ligere endene med biotinmerkede universal adaptere, osv. Deretter festes enden til underlaget slik som vist. 3) I neste trinn kutter man med IIS- og blunt end enzymet (i dette tilfellet illustrert med FokI og DraI). Dermed får man laget et nytt overheng i DNA biten som representerer de fire neste basene. 4) Dermed kan man tilsette konverteringsadaptere som konverterer disse basene til en signalkjede. DNA molekylene rettes så ut og avleses med f.eks. en fluorescensscanner. Ved avlesingen får man dermed informasjon om fire baser på grunnlag av posisjonen til DNA molekylene ene ende, og deretter informasjon om de fire neste på grunnlag av signalkjeden i den andre enden.

### Presortering av 8 basepar fordelt på 2 runder på en DNA chips

Det må videre presiseres at presorteringen selvfølgelig kan gjøres med færre eller flere permutasjoner enn 256 som illustrert i forrige avsnitt. Presorteringen kan også utføres i flere runder. Hvis man f.eks. lager DNA chips med 65.536 ulike ruter vil det bli mulig å identifisere 8 bp i presorteringen alene. For mange formål vil det holde for å kunne gjøre en vellykket rekonstruksjon. Presorteringen kan derfor fungere som en sekvenseringsmetode i seg selv, uten bruk av konverteringstrinn hvor man fester signalkjeder til mål DNA'en

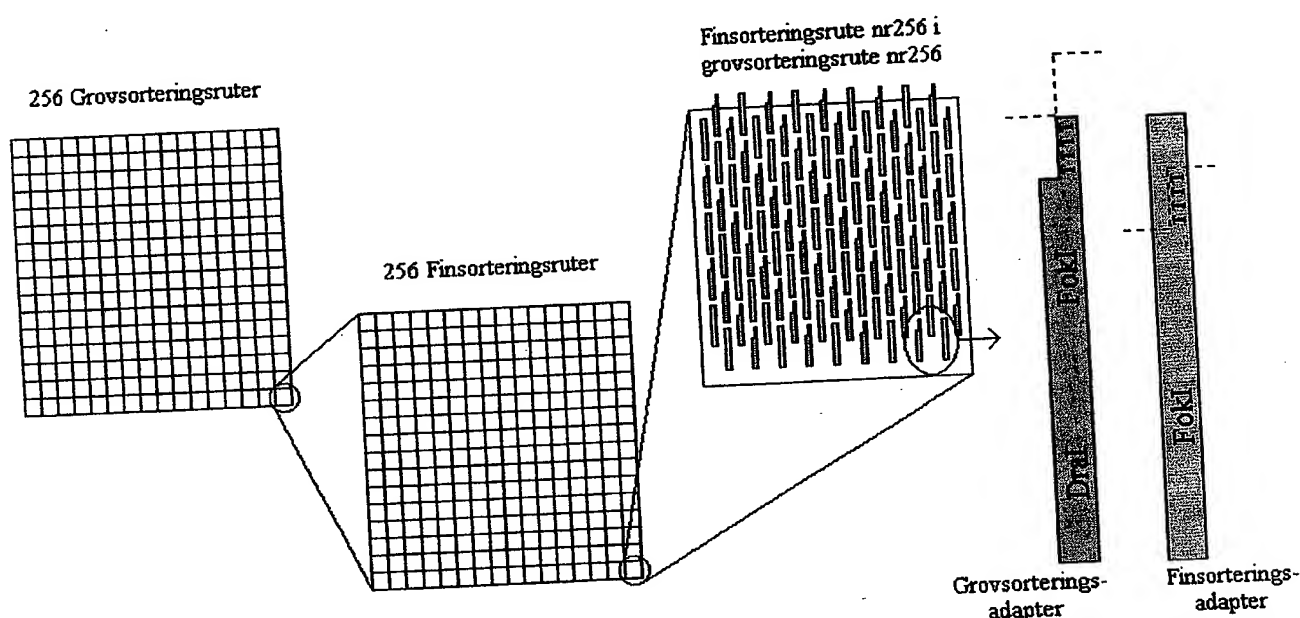


fig 33. Eksempel på DNA chips brukt i metode hvor presorteringen gjøres i to trinn. DNA chippen er grovinndelt i 256 ruter som er fysisk adskilt fra hverandre slik at væske ikke kan renne mellom rutene (ikke vist). Grovsorteringsrute 1 inneholder grovsorteringsadaptere med AAAA-overheng, rute 2 har AAAC-overheng, osv. I tillegg til overhenget har hver grovsorteringsadapter et bindingssete for en klasse IIS restriksjons endonuklease, f.eks. FokI. Ved bruk av den respektive endonukleasen vil det dannes et overheng, illustrert med stiplede linjer i de fire første baseparene i mål DNA'en som evt. ligger med overhenget. Hver av de 256 grovsorteringsrutene er videre fininndelt i 256 ruter hver seg, slik at avlesningsplaten totalt er inndelt i 65.536 finsorteringsruter. I hver finsorteringsrute er det finsorteringsadaptere som består av en ende med 4 variable bp samt et bindingssete for en klasse IIS restriksjons endonuklease, f.eks. FokI. Ved kutting med restriksjonsendonukleasen vil det dannes et overheng med utgangspunkt i de 4 variable bp slik som illustrert med stiplede linjer. Finsorteringsadapterene i finsorteringsrute 1 danner AAAA-overheng, i fininndelingsrute 2 dannes AAAC overheng, osv. Den samme fininndelingen finnes i alle grovsorteringsrutene.

Fig.33 illustrerer en DNA chip som kan brukes til en presortering av 8 bp i to trinn. Presorteringen innledes med å behandle DNA bitene slik at de danner overheng på 4bp. Det kan f.eks. gjøres ved å ligere DNA bitenes ender med adaptere inneholdende kuttsetter for en klasse IIS restriksjonsendonuklease som kutter de 4 første baseparene i DNA biten. Deretter fordeles løsningen med DNA biter på de 256 grovinndelingsrutene samtidig som det tilsettes ligase med egnet buffer. I grovsorteringsrute 1 vil DNA biter med TTTT overheng liggeres til grovsorteringsadapterene, i grovsorteringsrute 2 DNA biter med CTTT overheng, osv. Deretter vaskes rutene slik at DNA biter som ikke har ligert seg til adapterene skylles vekk. Deretter tilsettes den eller de aktuelle restriksjons

endonukleasene, f.eks. FokI med tilhørende buffer. Dermed kuttes DNA bitene vekk fra grovsorteringsadapterene idet det dannes et overheng bestående av de fire neste basene i DNA biten. Samtidig klargjøres finsorteringsadapterene ved at restriksjons endonukleasen lager et overheng med utgangspunkt i de fire variable basene. Når det i neste trinn tilsettes ligase vil DNA bitene finsorteres innenfor hver grovsorteringsrute. I finsorteringsrute 1 i grovsorteringsrute 1 vil det være festet DNA biter som begynte med AAAA AAAA mens det i finsorteringsrute 256 i grovsorteringsrute 256 vil være DNA biter som begynte med TTTT TTTT.

En potensiell kilde til feilsortering i det ovennevnte eksempelet er at grovsorteringsadapterene også vil kunne fungere som finsorteringsadaptere. Dette problemet kan imidlertid unngås ved at man har et kuttesete for en annen restriksjons endonuklease i grovsorteringsadapteren som gjør at man kan kutte vekk grovsorteringsadapterene før avlesningen. Selv om det ikke er nevnt i det ovenstående eksempelet, er det også viktig å terminere den enden av DNA biten som ikke festes til grovsorteringsadapterene. Det kan f.eks. gjøres med klenowutfylling.

De ulike prinsippene nevnt ovenfor kan selvfølgelig kombineres på svært mange måter. Man kan f.eks. gjøre de 256 første i reagensrør, deretter fordele rørene på 256 plater hvor man presorterer 256x256 bp slik som foreslått i fig... og til slutt konvertere 4 bp slik som foreslått i fig.... Ved avlesningen får man dermed informasjon om 24bp pluss deres omtrentlige lokalisasjon i mål sekvensen.

#### **Presortering basert på hybridisering**

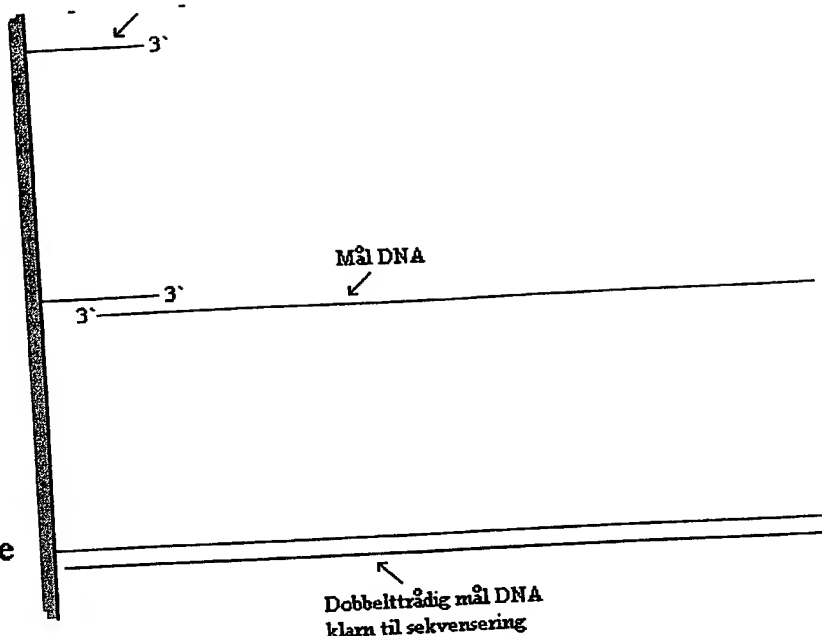
Et alternativ til presorteringsstrategiene ovenfor hvor man bruker ligeringer av overheng, er å basere presorteringen på ulike hybridiseringsreaksjoner. F.eks. kan man hybridisere enkelttrådig mål DNA til en DNA chips som består av adresser med alle permutasjoner av enkelttrådig oligonukleotider av en gitt lengde. Chipsen kan f.eks. være inndelt i 65.365 ruter som hver inneholder oligonukleotider på 8 baser. Adresse 1 kan f.eks. inneholde oligonukleotidet AAAAAAAA, adresse 2 AAAAAAAC, osv. Dermed inneholder chipsen alle permutasjoner av 8 baser.

Etter å ha hybridisert mål DNA'en rettes denne ut og merkes/ behandles slik at den 8 baser lange sekvensbiten på mål DNA'en som har hybridisert seg til DNA chipsen kan gis en posisjon i den totale mål sekvensen. Det kan f.eks. oppnås med "optical mapping". Hvis sekvensbitens posisjon kan bestemmes med en nøyaktighet på noen få kb eller bedre vil det som regel være mulig å foreta en vellykket rekonstruksjon av den totale mål sekvensen.

Hvis DNA bitene ved ovennevnte strategi er lange kan det være en fordel å forankre de hybridiserte DNA bitene til underlaget før utrettingen finner sted. En rekke teknikker er kjent i faget. F.eks. kan man bruke biotin/ streptavidin. Festemekanismene som er omtalt i patentsøknaden P986133; "Fremgangsmåte for DNA sekvensering" også brukes:

1) Hybridiser  
mål DNA

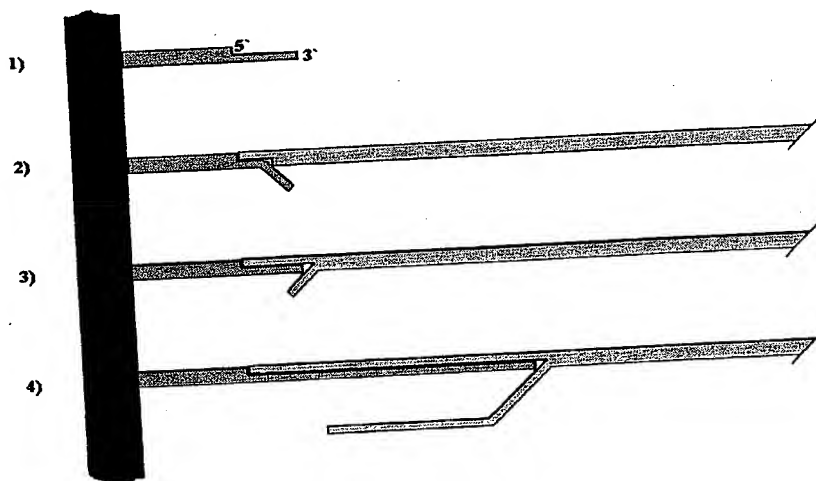
2) Polymerase  
ekstensjon



En rekke andre hybridiseringsstrategier kan selvfølgelig også tenkes. F.eks. kan man bruke PNA.

**Presortering basert på ligering pluss polymeraseekstensjon**

En annen strategi som både utnytter ligeringspesifisitet og polymerasespesifisitet kan være som nedenfor:



1) Utgangspunktet er presorteringsadaptere med overheng på f.eks. 8-10 baser. 2) En DNA bit med et overheng på 4 baser kan dermed ligeres til presorteringsadapteren hvis basene komplementerer de 4 innerste basene på presorteringsadapterens overheng. 3) Deretter utføres en polymerase ekstensjon. Forutsetningen for en vellykket polymeraseekstensjon er at resten av presorteringsadapterens overheng komplementerer DNA biten slik at den kan fungere som en primer. 4) En vellykket polymeraseekstensjon betyr med andre ord at DNA bitens ende komplementerer de 8-10 basene som utgjør presorteringsadapterens overheng på den aktuelle adressen.

**Andre presorteringsstrategier**

En rekke presorteringsstrategier kan tenkes. F.eks. kan samtlige av konverteringsprinsippene omtalt i patentsøknaden P986133; "Fremgangsmåte for DNA sekvensering" anvendes.

## **Generell definisjon av en sekvenseringsstrategi**

Med utgangspunkt i de ovenstående strategiene kan man bl.a. avlede følgende generelle sekvenseringsstrategi:

Sekvenseringsstrategi hvor fragmentert mål DNA (det vil si "DNA biter") posisjoneres til veldefinerte adresser på en DNA chips avhengig av en kort sekvensbit på den ene av mål DNA molekylens to ender. Mål DNA biter som ender med AAAAAAAAAA kan f.eks. festes til adresse 1, mens mål DNA biter som ender med TTTTTTTT kan f.eks. festes til adresse 65.536. Deretter behandles/ merkes på en slik måte at de korte sekvensbitene kan gis en mer eller mindre presis posisjon på mål sekvensen. Sistnevnte kan bl.a. oppnås med "optical mapping". Dermed vet man både hvilken sekvensbit som befinner seg på enden av et fragmentert mål DNA molekyl og hvor denne enden kan plasseres i den totale målsekvensen.

DNA bitene og mål DNA'en kan både være enkeltrådige, dobbelttrådige og trippeltrådige.

Den ovennevnte sekvenseringsstrategien kan videre kombineres med en strategi hvor en eller flere av baseparene konverteres til en signalkjede slik som omtalt i patentsøknaden P986133; "Fremgangsmåte for DNA sekvensering".

## **Reportergener, cis-regulatoriske elementer, o.l. som signaler**

Som tidligere nevnt åpner konverteringen til signalkjeder opp en rekke avlesningsmuligheter. I stedet for å bygge opp DNA kjeder som brukes som hybridiseringstemplater for fluorescerende prober, er det mulig å lage DNA kjeder som inneholder reportergener, cis-regulatoriske elementer, o.l. De konverterte DNA bitene kan deretter transformeres/ transfekteres inn i celler som så gjør informasjonen i DNA kjedene om til fysiske signaler som kan detekteres.

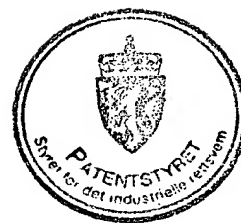




## P a t e n t k r a v

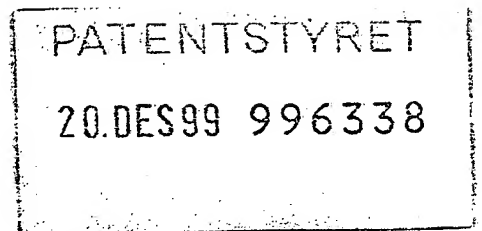
1.

Fremgangsmåter for sekvensanalyse, k a r a k t e r i s e r t v e d  
beskrivelsen.



1e

20 DES. 1999



Sammendrag

O. nr. E10515

Fremgangsmåter for sortering av DNA til bruk ved sekvensanalyse.

